Probabilistic Proofs of Existence of Rare Events

Noga Alon

Department of Mathematics
Sackler Faculty of Exact Sciences
Tel Aviv University
Ramat-Aviv, Tel Aviv 69978
ISRAEL

1. The Local Lemma

In a typical probabilistic proof of a combinatorial result, one usually has to show that the probability of a certain event is positive. However, many of these proofs actually give more and show that the probability of the event considered is not only positive but is large. In fact, most probabilistic proofs deal with events that hold with high probability, i.e., a probability that tends to 1 as the dimensions of the problem grow. For example, recall that a tournament on a set V of n players is a set of ordered pairs of distinct elements of V, such that for every two distinct elements x and y of V, either (x, y) or (y, x) is in the tournament, but not both. The name tournament is natural, since one can think on the set V as a set of players in which each pair participates in a single match, where (x, y) is in the tournament iff x defeated y. As shown by Erdös in [Er] for each $k \geq 1$ there are tournaments in which for every set of k players there is one who beats them all. The proof given in [Er] actually shows that for every fixed k if the number n of players is sufficiently large then almost all tournaments with n players satisfy this property, i.e., the probability that a random tournament with n players has the desired property tends to 1 as n tends to infinity.

On the other hand, there is a trivial case in which one can show that a certain event holds with positive, though very small, probability. Indeed, if we have n mutually independent events and each of them holds with probability at least $p > 0$, then the probability that all events hold simultaneously is at least p^n, which is positive, although it may be exponentially small in n.

It is natural to expect that the case of mutual independence can be generalized to that of rare dependencies, and provide a more general way of proving that certain events hold with positive, though small, proability. Such a generalization is, indeed, possible, and is stated in the following lemma, known as the Lovász Local Lemma. This simple lemma, first proved in [EL] is an extremely powerful tool, as it supplies a way for dealing with rare events.

Lemma 1.1 (The Local Lemma; General Case):

Let $A_1, A_2, \ldots, A_n$ be events in an arbitrary probability space. A directed graph $D = (V, E)$ on the set of vertices $V = \{1, 2, \ldots, n\}$ is called a *dependency digraph* for the events $A_1, \ldots, A_n$ if for each i, $1 \leq i \leq n$, the event A_i is mutually independent of all the events $\{A_j : (i, j) \notin E\}$. Suppose that $D = (V, E)$ is a dependency digraph for the above events and suppose there are real numbers $x_1, \ldots, x_n$ such that $0 \leq x_i < 1$ and $Pr(A_i) \leq x_i \prod_{(i,j)\in E}(1 - x_j)$ for all $1 \leq i \leq n$. Then $Pr\left(\bigwedge_{i=1}^{n} \overline{A}_j\right) \geq \prod_{i=1}^{n}(1 - x_i)$. In particular, with positive probability no event A_i holds.

Proof: We first prove, by induction on s, that for any $S \subset \{1, \ldots, n\}$, $|S| = s < n$ and any $i \notin S$

$$Pr\left(A_i | \bigwedge_{j \in S} \overline{A}_j\right) \leq x_i \ . \tag{1.1}$$

This is certainly true for $s = 0$. Assuming it holds for all $s' < s$, we prove it for S. Put $S_1 = \{j \in S; (i,j) \in E\}, S_2 = S \backslash S_1$. Then

$$Pr\left(A_i | \bigwedge_{j \in S} \overline{A}_j\right) = \frac{Pr\left(A_i \wedge (\bigwedge_{j \in S_1} \overline{A}_j | \bigwedge_{\ell \in S_2} \overline{A}_\ell\right)}{Pr\left(\bigwedge_{j \in S_1} \overline{A}_j | \bigwedge_{\ell \in S_2} \overline{A}_\ell\right)} \ . \tag{1.2}$$

To bound the numerator observe that since A_i is mutually independent of the events $\{A_\ell : \ell \in S_2\}$

$$Pr\left(A_i \wedge (\bigwedge_{j \in S_1} \overline{A}_j)| \bigwedge_{\ell \in S_2} \overline{A}_\ell\right) \leq Pr\left(A_i | \bigwedge_{\ell \in S_2} \overline{A}_\ell\right) = Pr(A_i) \leq x_i \prod_{(i,j) \in E} (1 - x_j) \ . \tag{1.3}$$

The denominator, on the other hand, can be bounded by the induction hypothesis. Indeed, suppose $S_1 = \{j_1, j_2, \ldots, j_r\}$. If $r = 0$ then the denominator is 1, and (1.1) follows. Otherwise

$$Pr\left(\overline{A}_{j_1} \wedge \overline{A}_{j_2} \wedge \cdots \wedge \overline{A}_{j_r} | \bigwedge_{\ell \in S_2} \overline{A}_\ell\right) = \left(1 - Pr(A_{j_1} | \bigwedge_{\ell \in S_2} \overline{A}_\ell)\right)$$

$$\cdot (1 - Pr\left(A_{j_2} | \overline{A}_{j_1} \wedge \bigwedge_{\ell \in S_2} \overline{A}_\ell\right) \cdot \ldots \cdot (1 - Pr\left(A_{j_r} | \overline{A}_{j_1} \wedge \ldots \wedge \overline{A}_{j_{r-1}} \wedge \bigwedge_{\ell \in S_2} \overline{A}_\ell\right)) \ . \tag{1.4}$$

$$\geq (1 - x_{j_1})(1 - x_{j_2}) \ldots (1 - x_{j_r}) \geq \prod_{(ij) \in E} (1 - x_j)$$

Substituting (1.3) and (1.4) into (1.2) we conclude that $Pr\left(A_i | \bigwedge_{j \in S} \overline{A}_j\right) \leq x_i$, completing the proof of the induction.

The assertion of Lemma 1.1 now follows easily, as

$$Pr\left(\bigwedge_{i=1}^{n} \overline{A}_i\right) = (1 - Pr(A_1)) \cdot (1 - Pr(A_2 | \overline{A}_1)) \cdot \ldots \cdot (1 - Pr(A_n | \bigwedge_{i=1}^{n-1} \overline{A}_i) \geq \prod_{i=1}^{n} (1 - x_i) \ ,$$

completing the proof. □

Corollary 1.2 (The Local Lemma; Symmetric Case): *Let $A_1, A_2, \ldots, A_n$ be events in an arbitrary probability space. Suppose that each event A_i is mutually independent of a set of all the other events A_j but at most d, and that $Pr(A_i) \le p$ for all $1 \le i \le n$. If*

$$ep(d+1) \le 1 \tag{1.5}$$

then $Pr\left(\bigwedge_{i=1}^{n} \overline{A}_i\right) > 0$.

Proof: If $d = 0$ the result is trivial. Otherwise, by the assumption there is a dependency digraph $D = (V, E)$ for the events $A_1, \ldots, A_n$ in which for each i, $\left|\{j : (i,j) \in E\}\right| \le d$. The result now follows from Lemma 1.1 by taking $x_i = 1/(d+1)(< 1)$ for all i and using the fact that for any $d \ge 2$ $\left(1 - \frac{1}{d+1}\right)^d > 1/e$. □

It is worth noting that as shown by Shearer in [Sh],the constant "e" is essentially best possible in inequality (1.5). Note also that the proof of Lemma 1.1 indicates that the conclusion remains true even when we replace the two assumptions that each A_i is mutually independent of $\{A_j : (i,j) \notin E)$ and that $Pr(A_i) \le x_i \prod_{(ij)\in E} (1 - x_j)$ by the weaker assumption that for each i and each $S_2 \subset \{1, \ldots, n\} \backslash \{j : (i,j) \in E\}$ $Pr\left(x_i \middle| \bigwedge_{j \in S_2} \overline{A}_j\right) \le x_i \prod_{(i,j)\in E} (1 - x_j)$. This turns out to be useful in certain applications.

In the next few sections we present various old and new applications of the Local Lemma for obtaining combinatorial results. There is no known proof of any of these results, which does not use the Local Lemma. It seems that the basic proof technique described here may be useful for many other combinatorial and non-combinatorial problems.

2. Property B and multicolored sets of real numbers

Recall that a hypergraph $H = (V, E)$ is simply a finite set V and a collection of subsets of it E. H has property B, (i.e. is 2-colorable), if there is a coloring of V by two colors so that no edge $f \in E$ is monochromatic.

Theorem 2.1. *Let $H = (V, E)$ be a hypergraph in which every edge has at least k elements, and suppose that each edge of H intersects at most d other edges. If $e(d+1) \le 2^{k-1}$ then H has property B.*

Proof: Color each vertex v of H, randomly and independently, either blue or red (with equal probability). For each edge $f \in E$, let A_f be the event that f is monochromatic. Clearly $Pr(A_f) = 2/2^{|f|} \le 1/2^{k-1}$. Moreover, each event A_f is clearly mutually independent of all the other events $A_{f'}$ for all edges f' that do not intersect f. The result now follows from Corollary 1.2. □

A special case of Theorem 2.1 is that for any $k \geq 9$, any k-uniform k-regular hypergraph H has property B. Indeed, since any edge f of such an H contains k vertices, each of which is incident with k edges (including f), it follows that f intersects at most $d = k(k-1)$ other edges. The desired result follows, since $e(k(k-1)+1) < 2^{k-1}$ for each $k \geq 9$. This special case has a different proof (see [AB]), which works for each $k \geq 8$. It seems, however, that in fact for each $k \geq 4$ each k-uniform k-regular hypergraph is 2-colorable.

The next result we consider, which appeared in the original paper of Erdös and Lovász, deals with k-colorings of the real numbers. For a k-coloring $c : I\!R \rightarrow \{1, 2, \ldots, k\}$ of the real numbers by the k colors $1, 2, \ldots, k$, and for a subset $T \subset I\!R$, we say that T is *multicolored* (with respect to c) if $c(T) = \{1, 2, \ldots, k\}$, i.e., if T contains elements of all colors.

Theorem 2.2. *Let m and k be two positive integers satisfying*

$$e\,(m(m-1)+1)\,k\left(1-\frac{1}{k}\right)^{m} \leq 1 \; . \tag{2.1}$$

Then, for any set S of m real numbers there is a k-coloring so that each translation $x + S$ (for $x \in I\!R$) is multicolored.

Notice that (2.1) holds whenever $m > (3 + o(1))k \log k$. There is no known proof of existence of any $m = m(k)$ with this property without using the local lemma.

Proof: We first fix a *finite* subset $X \subseteq I\!R$ and show the existence of a k-coloring so that each translation $x + S$ (for $x \in X$) is multicolored. This is an easy consequence of the Local Lemma. Indeed, put $Y = \bigcup_{x \in X}(x+S)$ and let $c : Y \rightarrow \{1, 2, \ldots, k\}$ be a random k-coloring of Y obtained by choosing, for each $y \in Y$, randomly and independently, $c(y) \in \{1, 2, \ldots, k\}$ according to a uniform distribution on $\{1, 2, \ldots, k\}$. For each $x \in X$, let A_x be the event that $x+S$ is not multicolored (with respect to c). Clearly $Pr(A_x) \leq k\left(1 - \frac{1}{k}\right)^m$. Moreover, each event A_x is mutually independent of all the other events $A_{x'}$ but those for which $(x+S) \cap (x'+S) \neq \emptyset$. As there are at most $m(m-1)$ such events the desired result follows from Corollary 1.2.

We can now prove the existence of a coloring of the set of all reals with the desired properties, by a standard compactness argument. Since the discrete space with k points is (trivially) compact, Tichonov's Theorem (which is equivalent to the axiom of choice) implies that an arbitrary product of such spaces is compact. In particular, the space of all functions from $I\!R$ to $\{1, 2, \ldots, k\}$, with the usual product topology, is compact. In this space for every fixed $x \in I\!R$, the set C_x of all colorings c, such that $x + S$ is multicolored is closed. (In fact, it is both open and closed, since a basis to the open sets is the set of all colorings whose values are prescribed in a finite number of places).

As we proved above, the intersection of any finite number of sets C_x is nonempty. It thus follows, by compactness, that the intersection of all sets C_x is nonempty. Any coloring in this intersection has the properties in the conclusion of Theorem 2.2. $\square$

Note that it is impossible, in general, to apply the Local Lemma to an infinite number of events and conclude that in some point of the probability space none of them holds. In fact, there are trivial examples of countably many mutually independent events A_i, satisfying $Pr(A_i) = 1/2$ and $\bigwedge_{i\geq 1} \overline{A}_i = \emptyset$. Thus the compactness argument is essential in the above proof.

3. Lower bounds for Ramsey numbers

The Ramsey number $R(k,l)$ is the minimum number n such that in any 2-coloring of the edges of the complete graph K_n on n vertices either there is a red K_k or a blue K_l. It is not too difficult to show that $R(k,l) \leq \binom{k+l-2}{k-1}$. The derivation of lower bounds for Ramsey numbers by Erdös in 1947 was one of the first applications of the probabilistic method. The Local Lemma provides a simple way of improving these bounds. Let us obtain, first, a lower bound for the diagonal Ramsey number $R(k,k)$. Consider a random 2-coloring of the edges of K_n. For each set S of k vertices of K_n, let A_s be the event that the complete graph on S is monochromatic. Clearly $Pr(A_S) = 2^{1-\binom{k}{2}}$. It is obvious that each event A_s is mutually independent of all the events A_T, but those which satisfy $|S \cap T| \geq 2$, since this is the only case in which the corresponding complete graphs share an edge. We can therefore apply Corollary 1.2 with $p = 2^{1-\binom{k}{2}}$ and $d = \binom{k}{2}\binom{n}{k-2}$ to conclude;

Proposition 3.1. If $e\left(\binom{k}{2}\binom{n}{k-2}+1\right)\cdot 2^{1-\binom{k}{2}} < 1$ then $R(k,k) > n$.

A short computation shows that this gives $R(k,k) > \frac{\sqrt{2}}{e}(1+o(1))\,k2^{k/2}$, only a small constant factor improvement on the bound obtained by the straightforward probabilistic method. Although this minor improvement is somewhat disappointing it is certainly not surprising; the Local Lemma is most powerful when the dependencies between events are rare, and this is not the case here. Indeed, there is a total number of $K = \binom{n}{k}$ events considered, and the maximum outdegree d in the dependency digraph is roughly $\binom{k}{2}\binom{n}{k-2}$. For large k and much larger n (which is the case of interest for us) we have $d > K^{1-O(1/k)}$, i.e., quite a lot of dependencies. On the other hand, if we consider small sets S, e.g., sets of size 3, we observe that out of the total $K = \binom{n}{3}$ of them each shares an edge with only $3(n-3) \approx K^{1/3}$. This suggests that the Local Lemma may be much more significant in improving the off-diagonal Ramsey numbers $R(k,\ell)$, especially if one of the parameters, say ℓ, is small. Let us consider, for example, following [Sp], the Ramsey number $R(k,3)$. Here, of course, we have to apply the nonsymmetric form of the Local Lemma. Let us 2-color the edges of K_n randomly and independently, where each edge is colored blue with probability p. For each set of 3 vertices

T, let A_T be the event that the triangle on T is blue. Similarly, for each set of k vertices S, let B_S be the event that the complete graph on S is red. Clearly $Pr(A_T) = p^3$ and $Pr(B_S) = (1-p)^{\binom{k}{2}}$. Construct a dependency digraph for the events A_T and B_S by joining two vertices by edges (in both directions) iff the corresponding complete graphs share an edge. Clearly, each A_T-node of the dependency graph is adjacent to $3(n-3) < 3n$ $A_{T'}$-nodes and to at most $\binom{n}{k}$ $B_{S'}$-nodes. Similarly, each B_S-node is adjacent to $\binom{k}{2}(n-k) < k^2n/2$ A_T nodes and to at most $\binom{n}{k}$ $B_{S'}$-nodes. It follows from the general case of the Local Lemma (Lemma 1.1) that if we can find a $0 < p < 1$ and two real numbers $0 \leq x < 1$ and $0 \leq y < 1$ such that

$$p^3 \leq x(1-x)^{3n}(1-y)^{\binom{n}{k}}$$

and

$$(1-p)^{\binom{k}{2}} \leq y(1-x)^{k^2n/2}(1-y)^{\binom{n}{k}}$$

then $R(k,3) > n$.

Our objective is to find the largest possible $k = k(n)$ for which there is such a choice of p, x and y. An elementary (but tedious) computation shows that the best choice is when $p = c_1 n^{-1/2}$, $k = c_2 n^{1/2} \log n, x = c_3/n^{3/2}$ and $y = \frac{c_4}{e^{n^{1/2}\log^2 n}}$. This gives that $R(k,3) > c_5 k^2/\log^2 k$. A similar argument gives that $R(k,4) > k^{5/2+o(1)}$. In both cases the amount of computation required is considerable. However, the hard work does pay (in this case, at least); the bound $R(k,3) > c_5 k^2/\log^2 k$ matches a lower bound of Erdös , proved in 1961, obtained by a highly complicated probabilistic argument. The bound above for $R(k,4)$ is better than any bound for $R(k,4)$ known to be proved without the Local Lemma.

4. A geometric result

A family of open unit balls F in the 3-dimensional Euclidean space $\mathbb{R}^3$ is called a *k-fold covering* of $\mathbb{R}^3$ if any point $x \in \mathbb{R}^3$ belongs to at least k balls. In particular, a 1-fold covering is simply called a *covering*. A k-fold covering $\mathcal{F}$ is called *decomposable* if there is a partition of $\mathcal{F}$ into two pairwise disjoint families $\mathcal{F}_1$ and $\mathcal{F}_2$, each being a covering of $\mathbb{R}^3$. Mani and Pach ([MP]) constructed, for any integer $k \geq 1$, a non-decomposable k-fold covering of $\mathbb{R}^3$ by open unit balls. On the other hand they proved that any k-fold covering of $\mathbb{R}^3$ in which no point is covered by more than $c2^{k/3}$ balls is decomposable. This reveals a somewhat surprising phenomenon that it is more difficult to decompose coverings that cover some of the points of $\mathbb{R}^3$ too often, than to decompose coverings that cover every point about the same number of times. The exact statement of the Mani-Pach Theorem is the following.

Theorem 4.1. Let $\mathcal{F} = \{B_i\}_{i\in I}$ be a k-fold covering of the 3 dimensional Euclidean space by open unit balls. Suppose, further, than no point of $I\!R^3$ is contained in more than t members of $\mathcal{F}$. If

$$e \cdot t^3 2^{18}/2^{k-1} \leq 1$$

then $\mathcal{F}$ is decomposable.

Proof: Let $\{C_j\}_{j\in J}$ be the connected components of the set obtained from $I\!R^3$ by deleting all the boundaries of the balls B_i in $\mathcal{F}$. Let $H = (V(H), E(H))$ be the (infinite) hypergraph defined as follows; the set of vertices of H, $V(H)$ is simply $\mathcal{F} = \{B_i\}_{i\in I}$. The set of edges of H is $E(H) = \{E_j\}_{j\in J}$, where $E_j = \{B_i : i \in I \text{ and } C_j \subseteq B_i\}$. Since $\mathcal{F}$ is a k-fold covering, each edge E_j of H contains at least k vertices. We claim that each edge of H intersects less than $t^3 2^{18}$ other edges of H. To prove this claim, fix an edge E_ℓ, corresponding to the connected component C_ℓ, where $\ell \in J$. Let E_j be an arbitrary edge of H, corresponding to the component C_j, that intersects E_ℓ. Then there is a ball B_i containing both C_ℓ and C_j. Therefore, any ball that contains C_j intersects B_i. It follows that all the unit balls that contain or touch a C_j, for some j that satisfies $E_j \cap E_\ell \neq \emptyset$ are contained in a ball B of radius 4. As no point of this ball is covered more than t times we conclude, by a simple volume argument, that the total number of these unit balls is at most $t \cdot 4^3 = t \cdot 2^6$. It is not too difficult to check that m balls in $I\!R^3$ cut $I\!R^3$ into less than m^3 connected components, and since each of the above C_j is such a component we have $\left|\{j : E_j \cap E_\ell \neq \emptyset\}\right| < (t \cdot 2^6)^3 = t^3 2^{18}$, as claimed.

Consider, now, any finite subhypergraph L of H. Each edge of L has at least k vertices, and it intersects at most $d < t^3 2^{18}$ other edges of L. Since, by assumption, $e(d+1) \leq 2^{k-1}$, Theorem 2.1 (which is a simple corollary of the local lemma), implies that L is 2-colorable. This means that one can color the vertices of L blue and red so that no edge of L is monochromatic. Since this holds for any finite L, a compactness argument, analogous to the one used in the proof of Theorem 2.2, shows that H is 2-colorable. Given a 2-coloring of H with no monochromatic edges, we simply let $\mathcal{F}_1$ be the set of all blue balls, and $\mathcal{F}_2$ be the set of all red ones. Clearly, each $\mathcal{F}_i$ is a covering of $I\!R^3$, completing the proof of the theorem. □

It is worth noting that Theorem 4.1 can be easily generalized to higher dimensions. We omit the detailed statement of this generalization.

5. The Linear Arboricity of Graphs

A *linear forest* is a forest (i.e., an acyclic simple graph) in which every connected component is a path. The *linear arboricity* $la(G)$ of a graph G is the minimum number of linear forests in

G, whose union is the set of all edges of G. This notion was introduced by Harary as one of the covering invariants of graphs. The following conjecture, known as the *linear arboricity conjecture*, was raised in [AEH]:

Conjecture 5.1. (The linear arboricity conjecture). The linear arboricity of every d-regular graph is $\lceil (d+1)/2 \rceil$.

Notice that since every d-regular graph G on n vertices has $nd/2$ edges, and every linear forest in it has at most $n-1$ edges, the inequality

$$\mathrm{la}(G) \geq \frac{nd}{2(n-1)} > \frac{d}{2}$$

is immediate. Since $\mathrm{la}(G)$ is an integer this gives $\mathrm{la}(G) \geq \lceil (d+1)/2 \rceil$. The difficulty in Conjecture 5.1 lies in proving the converse inequality: $\mathrm{la}(G) \leq \lceil (d+1)/2 \rceil$. Note also that since every graph G with maximum degree Δ is a subgraph of a Δ-regular graph (which may have more vertices, as well as more edges than G), the linear arboricity conjecture is equivalent to the statement that the linear arboricity of every graph G with maximum degree Δ is at most $\lceil (\Delta+1)/2 \rceil$.

Although this conjecture received a considerable amount of attention, the best general result concerning it, proved without any probabilistic arguments, is that $la(G) \leq \lceil 3\Delta/5 \rceil$ for even Δ and that $la(G) \leq \lceil (3\Delta+2)/5 \rceil$ for odd Δ. In this section we prove that for every $\varepsilon > 0$ there is a $\Delta_0 = \Delta_0(\varepsilon)$ such that for every $\Delta \geq \Delta_0$, the linear arboricity of every graph with maximum degree Δ is less than $\left(\frac{1}{2}+\varepsilon\right)\Delta$. This result appears in [Al] and its proof relies heavily on the local lemma. Here we present a simpler proof , that supplies a better estimate for the error term. This proof requires certain preparations, some of which are of independent interest. It is convenient to deduce the result for undirected graphs from its directed version.

A *d-regular digraph* is a directed graph in which the indegree and the outdegree of every vertex is precisely d. A linear directed forest is a directed graph in which every connected component is a directed path. The *di-linear arboricity* $\mathrm{dla}(G)$ of a directed graph G is the minimum number of linear directed forests in G whose union covers all edges of G. The directed version of the Linear Arboricity Conjecture, first stated in [NP] is;

Conjecture 5.2. For every d-regular digraph D,

$$\mathrm{dla}(D) = d+1 \ .$$

Note that since the edges of any (connected) undirected $2d$-regular graph G can be oriented along an Euler cycle, so that the resulting oriented digraph is d-regular, the validity of Conjecture 5.2 for d implies that of Conjecture 5.1 for $2d$.

It is easy to prove that any graph with n vertices and maximum degree d contains an independent set of size at least $n/(d+1)$. The following proposition shows that at the price of decreasing the size of such a set by a constant factor we can guarantee that it has a certain structure.

Proposition 5.3. *Let $H = (V, E)$ be a graph with maximum degree d, and let $V = V_1 \cup V_2 \cup \cdots \cup V_r$ be a partition of V into r pairwise disjoint sets. Suppose each set V_i is of cardinality $|V_i| \geq 25d$. Then there is an independent set of vertices $W \subseteq V$, that contains at least one vertex from each V_i.*

Proof: Clearly we may assume that each set V_i is of cardinality precisely $g = 25d$ (otherwise, simply replace each V_i by a subset of cardinality g of it, and replace H by its induced subgraph on the union of these r new sets). Put $p = 1/25d$, and let us pick each vertex of H, randomly and independently, with probability p. Let W be the random set of all vertices picked. To complete the proof we show that with positive probability W is an independent set of vertices that contains a point from each V_i. For each i, $1 \leq i \leq r$, let S_i be the event that $W \cap V_i = \emptyset$. Clearly $Pr(S_i) = (1-p)^g$. For each edge f of H, let A_f be the event that W contains both ends of f. Clearly, $Pr(A_f) = p^2$. Moreover, each event S_i is mutually independent of all the events

$$\{S_j : 1 \leq j \leq r, j \neq i\} \cup \{A_f : f \cap V_i = \emptyset\} \ .$$

Similarly, each event, A_f is mutually independent of all the events

$$\{S_j : S_j \cap f = \emptyset\} \cap \{A_{f'} : f' \cap f = \emptyset\} \ .$$

Therefore, there is a dependency digraph for the events $\{S_i : 1 \leq i \leq r\} \cup \{A_f : f \in E\}$ in which each S_j-node is adjacent to at most $g \cdot d$ A_f-nodes (and to no $S_{j'}$-nodes), and each A_f-node is adjacent to at most 2 S_j-nodes, and at most $2d-2$ $A_{f'}$-nodes. It follows from Lemma 1.1 (the local lemma) that if we can find two numbers x and $y, 0 \leq x < 1, 0 \leq y < 1$ so that

$$(1-p)^g = \left(1 - \frac{1}{25d}\right)^{25d} = Pr(S_i) < x(1-y)^{gd} = x(1-y)^{25d^2} \tag{5.1}$$

and

$$p^2 = \frac{1}{(25d)^2} = Pr(A_f) < y(1-y)^{2d-2} \cdot (1-x)^2 \tag{5.2}$$

then $Pr(\bigwedge_{f \in E} \overline{A}_f \bigwedge_{1 \leq i \leq r} \overline{S}_i) > 0$. One can easily check that $x = \frac{1}{2}, y = 1/100d^2$ satisfy (5.1) and (5.2). Indeed

$$\left(\frac{1}{2}\right)\left(1 - \frac{1}{100d^2}\right)^{25d^2} \geq \frac{1}{2}\left(1 - \frac{25d^2}{100d^2}\right) = \frac{3}{8} \geq \frac{1}{e} \geq \left(1 - \frac{1}{25d}\right)^{25d}$$

and

$$\frac{1}{100d^2}\left(1-\frac{1}{100d^2}\right)^{2d-2}\left(\frac{1}{2}\right)^2 \geq \frac{1}{400d^2}\left(1-\frac{1}{50d}\right) > \frac{1}{(25d)^2}$$

Therefore,

$$Pr\left(\bigwedge_{f\in E}\overline{A}_f \bigwedge_{1\leq i\leq r}\overline{S}_i\right) > 0\ ,$$

i.e., with positive probability, none of the events S_i or A_f hold for W. In particular, there is at least one choice for such $W \subseteq V$. But this means that this W is an independent set, containing at least one vertex from each V_i. This completes the proof. □

Proposition 5.3 suffices to proves Conjecture 5.2 for digraphs with no short directed cycle. Recall that the directed girth of a digraph is the minimum length of a directed cycle in it.

Theorem 5.4. *Let $G = (U, F)$ be a d-regular digraph with directed girth $g \geq 50d$. Then*

$$\mathrm{dla}(G) = d + 1\ .$$

Proof: As is well known, F can be partitioned into d pairwise disjoint 1-regular spanning subgraphs $F_1, \ldots, F_d$ of G. (This is an easy consequence of the Hall-König Theorem; let H be the bipartite graph whose two classes of vertices A and B are copies of U, in which $u \in A$ is joined to $v \in B$ iff $(u, v) \in F$. Since H is d-regular its edges can be decomposed into d perfect matchings, which correspond to d 1-regular spanning subgraphs of G.) Each F_i is a union of vertex disjoint directed cycles $C_{i_1}, C_{i_2}, \ldots, C_{ir_i}$. Let $V_1,\ V_2, \ldots, V_r$ be the sets of *edges* of all the cycles $\{C_{ij} : 1 \leq i \leq d, 1 \leq j \leq r_i\}$. Clearly $V_1, V_2, \ldots, V_r$ is a partition of the set F of all edges of G, and by the girth condition, $|V_i| \geq g \geq 50d$ for all $1 \leq i \leq r$. Let H be the line graph of G, i.e., the graph whose set of vertices is the set F of edges of G in which two edges are adjacent iff they share a common vertex in G. Clearly H is $2d - 2$ regular. As the cardinality of each V_i is at least $50d \geq 25(2d - 2)$, there is, by Proposition 5.3, an independent set of H containing a member from each V_i. But this means that there is a matching M in G, containing at least one edge from each cycle C_{ij} of the 1-factors $F_1, \ldots, F_d$. Therefore $M, F_1\backslash M, F_2\backslash M, \ldots, F_d\backslash M$ are $d+1$-directed forests in G (one of which is a matching) that cover all its edges. Hence

$$\mathrm{dla}(G) \leq d + 1\ .$$

As G has $|U| \cdot d$ edges and each directed linear forest can have at most $|U| - 1$ edges,

$$\mathrm{dla}(G) \geq |U|d/(|U| - 1) > d\ .$$

Thus $\mathrm{dla}(G) = d + 1$, completing the proof. $\square$

The last theorem shows that the assertion of Conjecture 5.2 holds for digraphs with sufficiently large (directed) girth. In order to deal with digraphs with small girth, we show that most of the edges of each regular digraph can be decomposed to a relatively small number of almost regular digraphs with high girth. To do this, we need the following statement, which is proved using the local lemma.

Lemma 5.5. *Let $G = (V, E)$ be a d-regular directed graph, where $d \geq 100$, and let p be an integer satisfying $10\sqrt{d} \leq p \leq 20\sqrt{d}$. Then, there is a p-coloring of the vertices of G by the colors $0, 1, 2, \ldots, p-1$ with the following property; for each vertex $v \in V$ and each color i, the numbers $N^+(v, i) = \left|\{u \in V; (v, u) \in E \text{ and } u \text{ is colored } i\}\right|$ and $N^-(v, i) = \left|\{u \in V : (u, v) \in E \text{ and } u \text{ is colored } i\}\right|$ satisfy:*

$$\begin{aligned} \left|N^+(v,i) - \tfrac{d}{p}\right| &\leq 3\sqrt{d/p}\sqrt{\log d}\ , \\ \left|N^-(v,i) - \tfrac{d}{p}\right| &\leq 3\sqrt{d/p}\sqrt{\log d}\ . \end{aligned} \tag{5.3}$$

Proof: Let $f : V \to \{0, 1, \ldots, p-1\}$ be a random vertex coloring of V by p colors, where for each $v \in V$, $f(v) \in \{0, 1, \ldots, p-1\}$ is chosen according to a uniform distribution. For every vertex $v \in V$ and every color $i, 0 \leq i < p$, let $A^+_{v,i}$ be the event that the number $N^+(v, i)$ of neighbors of v in G whose color is i does not satisfy inequality (5.3). Clearly, $N^+(v, i)$ is a Binomial random variable with expectation $\frac{d}{p}$ and standard deviation $\sqrt{\frac{d}{p}(1 - \frac{1}{p})} < \sqrt{\frac{d}{p}}$. Hence, by the standard estimates for Binomial distributions, for every $v \in V$ and $0 \leq i < p$

$$Pr(A^+_{v,i}) < e^{-\frac{9 \log d}{2}} < 1/d^4\ .$$

Similarly, if $A^-_{v,i}$ is the event that the number $N^-(v, i)$ violates (5.3) then

$$Pr(A^-_{v,i}) < 1/d^4\ .$$

Clearly, each of the events $A^+_{v,i}$ or $A^-_{v,i}$ is mutually independent of all the events $A^+_{u,j}$ or $A^-_{u,j}$ for all vertices $u \in V$ that do not have a common neighbor with v in G. Therefore, there is a dependency digraph for all our events with maximum degree $\leq 2d^2 \cdot p$. Since $e \cdot \frac{1}{d^4}(2d^2p + 1) < 1$, Corollary 1.2, (i.e., the symmetric form of the Local Lemma), implies that with positive probability no event $A^+_{v,i}$ or $A^-_{v,i}$ occurs. Hence, there is a coloring f which satisfies (5.3) for all $v \in V$ and $0 \leq i < p$, completing the proof. $\square$

We are now ready to deal with general regular digraphs. Let $G = (V, E)$ be an arbitrary d-regular digraph. Throughout the argument we assume, whenever it is needed, that d is sufficiently large. Let p be a prime satisfying $10d^{1/2} \leq p \leq 20d^{1/2}$ (it is well known that for every n there is a prime between n and $2n$). By Lemma 5.5 there is a vertex coloring $f : V \to \{0, 1, \ldots, p-1\}$ satisfying (5.3). For each i, $0 \leq i < p$, let $G_i = (V, E_i)$ be the spanning subdigraph of G defined by $E_i = \{(u,v) \in E : f(v) \equiv (f(u) + i) \bmod p\}$. By inequality (5.3) the maximum indegree Δ_i^- and the maximum outdegree Δ_i^+ in each G_i is at most $\frac{d}{p} + 3\sqrt{\frac{d}{p}}\sqrt{\log d}$. Moreover, for each $i > 0$, the length of every directed cycle in G_i is divisible by p. Thus, the directed girth g_i of G_i is at least p. Since each G_i can be completed, by adding vertices and edges, to a Δ_i-regular digraph with the same girth g_i and with $\Delta_i = \max\,(\Delta_i^+, \Delta_i^-)$, and since $g_i > 50\Delta_i$ (for all sufficiently large d), we conclude, by Theorem 5.4, that $\mathrm{dla}(G_i) \leq \Delta_i + 1 \leq \frac{d}{p} + 3\sqrt{\frac{d}{p}}\sqrt{\log d} + 1$ for all $1 \leq i < p$. For G_0, we only apply the trivial inequality

$$\mathrm{dla}(G_0) \leq 2\Delta_0 \leq 2\frac{d}{p} + 6\sqrt{\frac{d}{p}}\sqrt{\log d}$$

obtained by, e.g., embedding G_0 as a subgraph of a Δ_0-regular graph, splitting the edges of this graph into Δ_0 1-regular spanning subgraphs, and breaking each of these 1-regular spanning subgraphs into two linear directed forests. The last two inequalities, together with the fact that $10\sqrt{d} \leq p \leq 20\sqrt{d}$ imply

$$\mathrm{dla}(G) \leq d + \frac{d}{p} + 3\sqrt{pd}\sqrt{\log d} + 3\sqrt{\frac{d}{p}}\sqrt{\log d} + p - 1 \leq d + c \cdot d^{3/4}(\log d)^{1/2}$$

We have thus proved;

Theorem 5.6. *There is an absolute constant $c > 0$ such that for every d-regular digraph G*

$$\mathrm{dla}(G) \leq d + cd^{3/4}(\log d)^{1/2}\ . \square$$

We note that by being a little more careful, we can improve the error term to $c\ d^{2/3}(\log d)^{1/3}$. Since the edges of any undirected $d = 2f$-regular graph can be oriented so that the resulting digraph is f-regular, and since any $(2f-1)$-regular undirected graph is a subgraph of a $2f$-regular graph the last theorem implies;

Theorem 5.7. *There is an absolute constant $c > 0$ such that for every undirected d-regular graph G*

$$\mathrm{la}(G) \leq \frac{d}{2} + cd^{3/4}(\log d)^{1/2}\ . \qquad \square$$

6. Latin Transversals

Following the proof of the local lemma we noted that the mutual independency assumption in this lemma can be replaced by the weaker assumption that the conditional probability of each event, given the mutual non-occurrence of an arbitrary set of events, each nonadjacent to it in the dependency digraph, is sufficiently small. In this section we describe an application, from [ES], of this modified version of the lemma. Let $A = (a_{ij})$ be an n by n matrix with, say, integer entries. A permutation π is called a *Latin transversal* (of A) if the entries $a_{i\pi(i)}$ $(1 \le i \le n)$ are all distinct.

Theorem 6.1. *Suppose $k \le (n-1)/(4e)$ and suppose that no integer appears in more than k entries of A. Then A has a Latin Transversal.*

Proof: Let π be a random permutation of $\{1, 2, \ldots, n\}$, chosen according to a uniform distribution among all possible $n!$ permutations. Denote by T the set of all ordered fourtuples (i, j, i', j') satisfying $i < i', j \ne j'$ and $a_{ij} = a_{i'j'}$. For each $(i, j, i', j') \in T$, let $A_{iji'j'}$ denote the event that $\pi(i) = j$ and $\pi(i') = j'$. The existence of a Latin transversal is equivalent to the statement that with positive probability none of these events hold. Let us define a symmetric digraph, (i.e., a graph) G on the vertex set T by making (i, j, i', j') adjacent to (p, q, p', q') if and only if $\{i, j'\} \cap \{p, p'\} \ne \emptyset$ or $\{j, j'\} \cap \{q, q'\} \ne \emptyset$. Thus, these two fourtuples are not adjacent iff the four cells $(i, j), (i', j'), (p, q)$ and (p', q') occupy four distinct rows and columns of A. The maximum degree of G is less than $4nk$; indeed, for a given $(i, j, i', j') \in T$ there are $4n$ choices of (p, q) with either $p \in \{i, i'\}$ or $q \in \{j, j'\}$, and for each of these choices of (p, q) there are less than k choices for $(p', q') \ne (p, q)$ with $a_{pq} = a_{p'q'}$. Since $e \cdot 4nk \cdot \frac{1}{n(n-1)} \le 1$, the desired result follows from the above mentioned strengthening of the symmetric version of the Local Lemma, if we can show that

$$Pr(A_{iji'j'} \,|\, \bigwedge_S \overline{A}_{pqp'q'}) \le 1/n(n-1) \tag{6.1}$$

for any $(i, j, i'j') \in T$ and any set S of members of T which are nonadjacent in G to $(i, j, i'j')$. By symmetry, we may assume that $i = j = 1, i' = j' = 2$ and that hence none of the p's nor q's are either 1 or 2. Let us call a permutation π *good* if it satisfies $\bigwedge_S \overline{A}_{pqp'q'}$, and let S_{ij} denote the set of all good permutations π satisfying $\pi(1) = i$ and $\pi(2) = j$. We claim that $|S_{12}| \le |S_{ij}|$ for alll $i \ne j$. Indeed, suppose first that $i, j > 2$. For each good $\pi \in S_{12}$ define a permutation π^* as follows. Suppose $\pi(x) = i$ $\pi(y) = j$. Then define $\pi^*(1) = i, \pi^*(2) = j, \pi^*(x) = 1, \pi^*(y) = 2$ and $\pi^*(t) = \pi(t)$ for all $t \ne 1, 2, x, y$. One can easily check that π^* is good, since the cells $(1, i), (2, j), (x, 1), (y, 2)$ are not part of any $(p, q, p', q') \in S$. Thus $\pi^* \in S_{ij}$, and since the mapping $\pi \to \pi^*$ is injective

$\left|S_{12}\right| \leq \left|S_{ij}\right|$, as claimed. Similarly one can define injective mappings showing that $\left|S_{12}\right| \leq \left|S_{ij}\right|$ even when $\{i,j\} \cap \{1,2\} \neq \emptyset$. It follows that $Pr(A_{1122} \bigwedge \bigwedge_S \overline{A}_{pqp'q'}) \leq Pr(A_{1i2j} \bigwedge \bigwedge_S \overline{A}_{pqp'q'})$ for all $i \neq j$ and hence that $Pr(A_{1122} \big| \bigwedge \bigwedge_S \overline{A}_{pqp'q'}) \leq 1/n(n-1)$. By symmetry, this implies (6.1) and completes the proof. □

7. Cycles in Directed Graphs.

The last example we consider is an extremely simple, yet surprising, application of the local lemma. Let $D = (V, E)$ be a simple directed graph with minimum outdegree δ and maximum indegree Δ.

Theorem [AL]. *If* $e(\Delta\delta + 1)\left(1 - \frac{1}{k}\right)^{\delta} < 1$ *then* D *contains a (directed, simple) cycle of length* $0 (\mathrm{mod}\, k)$.

Proof: Clearly we may assume that every outdegree is precisely δ, since otherwise we can consider a subgraph of D with this property.

Let $f : V \to \{0, 1, \ldots, k-1\}$ be a random coloring of V, obtained by choosing, for each $v \in V, f(v) \in \{0, \ldots, k-1\}$ independently, according to a uniform distribution. For each $v \in V$, let A_v denote the event that there is no $u \in V$, with $(v, u) \in E$ and $f(u) \equiv (f(v) + 1) (\mathrm{mod}\, k)$. Clearly $Pr(A_v) = \left(1 - \frac{1}{k}\right)^{\delta}$. One can easily check that each event A_v is mutually independent of all the events A_u but those satisfying

$$N^+(v) \cap \left(u \bigcup N^+(u)\right) \neq \emptyset \ ,$$

where here $N^+(v) = \{w \in V : (v, w) \in E\}$. The number of such u's is at most $\Delta\delta$ and hence, by our assumption and by the Local Lemma, (Corollary 1.2), $Pr\left(\bigwedge_{v \in V} \overline{A}_v\right) > 0$. Therefore, there is an $f : V \to \{0, 1, \ldots, k-1\}$ such that for every $v \in V$ there is a $u \in V$ with $(*)(v, u) \in E$ and $f(u) \equiv (f(v) + 1) (\mathrm{mod}\, k)$. Starting at an arbitrary $v = v_0 \in V$ and applying $(*)$ repeatedly we obtain a sequence $v_0, v_1, v_2, \ldots$ of vertices of D so that $(v_i, v_{i+1}) \in E$ and $f(v_{i+1}) \equiv (f(v_i) + 1) (\mathrm{mod}\, k)$ for all $i \geq 0$. Let j be the minimum integer so that there is an $\ell < j$ with $v_\ell = v_j$. The cycle $v_\ell v_{\ell+1} v_{\ell+2} \cdots v_j = v_\ell$ is a directed simple cycle of D whose length is divisible by k. □

8. The Algorithmic Aspect.

When the probabilistic method is applied to prove that a certain event holds with high probability, it often supplies an efficient deterministic, or at least randomized, algorithm for the corresponding problem.

By applying the Local Lemma we often manage to prove that a given event holds with positive probability, although this probability may be exponentially small in the dimensions of the problem. Consequently, these proofs usually provide no polynomial algorithms for the corresponding problems. To be specific; in Section 2 we showed that any 9-regular 9-uniform hypergraph $H = (V, E)$ is 2-colorable. Can we actually find a legal 2-coloring of H in polynomial, or expected polynomial (in $|V| + |E|$) time? Similarly, can we find, efficiently, a Latin transversal in an n by n matrix in which no entry appears more than $n/15$ times? As shown in the previous section any 10-regular digraph contains an even directed simple cycle. Can we find efficiently such a cycle? There are no known efficient algorithms to any of these, and the problem of finding more effective versions of all these proofs remains an open and intriguing challenge.

References

[AEH] J. Akiyama, G. Exoo and F. Harary, Covering and packing in graphs IV: Linear arboricity, Networks 11(1981), 69–72.

[Al] N. Alon, The linear arboricity of graphs, Israel J. Math. 62(1988), 311–325.

[AB] N. Alon and Z. Bregman, Every 8-uniform 8-regular hypergraph is 2-colorable, Graphs and Combinatorics 4(1988), 303-305.

[AL] N. Alon and N. Linial, Cycles of length 0 modulo k in directed graphs, J. Combinatorial Theory, Series B, in press.

[Er] P. Erdös , On a problem in Graph Theory, Math. Gazette 47(1963), 220–223.

[EL] P. Erdös and L. Lovász, Problems and results on 3-chromatic hypergraphs and some related questions, in Infinite and Finite Sets (A. Hajnal et al., eds.), North-Holland, Amsterdam, 1975, pp. 609–628.

[ES] P. Erdös and J. Spencer, Lopsided Lovász Local Lemma and Latin Transversals, to appear.

[MP] P. Mani-Levitska and J. Pach, Decomposition problems for multiple coverings with unit balls, to appear.

[NP] A. Nakayama and B. Peroche, Linear arboricity of digraphs, Networks 17(1987), 39–53.

[Sh] J. Shearer, On a problem of Spencer, Combinatorica 5 (1985), 241–245.

[Sp] J. Spencer, Asymptotic lower bounds for Ramsey functions, Discrete Math. 20(1977), 69–76.